

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 6, 2002, 19:53:16 : Search time 2690 Seconds
(without alignments)
16498.803 Million cell updates/sec
Title: US-10-025-514-15
Perfect score: 1525
Sequence: 1 tctagaccatgaagaccct.....ccagtcaaggcctagtcgac 1525

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	GenEmbl :	Database				Description
		1	2	3	4	
		1: gb_ba.*	2: gb_htg.*	3: gb_in.*	4: gb_om.*	
		5: gb_ov.*	6: gb_pat.*	7: gb_ph.*	8: gb_pl.*	
		9: gb_pr.*	10: gb_ro.*	11: gb_sts.*	12: gb_sy.*	
		13: gb_un.*	14: gb_vl.*	15: em_ba.*	16: em_fun.*	
		17: em_hum.*	18: em_in.*	19: em_mu.*	20: em_on.*	
		21: em_or.*	22: em_ov.*	23: em_pat.*	24: em_ph.*	
		25: em_pl.*	26: em_ro.*	27: em_sts.*	28: em_un.*	
		29: em_vl.*	30: em_htg_hum.*	31: em_htg_inv.*	32: em_htg_other.*	
		33: em_htg_mus.*	34: em_htg_pln.*	35: em_htg_fod.*	36: em_htg_mam.*	
		37: em_htg_vrt.*	38: em_sv.*	39: em_htgo_hum.*	40: em_htgo_mus.*	
		41: em_htgo_other.*				

ALIGNMENTS			
RESULT 1	AR111412	1308 bp	DNA
LOCUS	Sequence 6 from patent US 6127145.	linear	PAT 14-FEB-2001
DEFINITION	AR111412		
ACCESSION	AR111412		
VERSION	AR111412.1	GI:12828260	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1308)		
AUTHORS	Sutliff, T.D. and Rodriguez, R.L.		
TITLE	Production of alpha-subunit 1 - antitrypsin in plants		
JOURNAL	Patent: US 6127145-A 6 03-OCT-2000;		
FEATURES	Location/Qualifiers		

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
1	629.4	41.3	1308	6	AR111412	AR111412 Sequence	
2	628.4	41.2	1185	6	AR111411	AR111411 Sequence	
3	437	28.7	1345	9	HUMALATB	M11465 Human alpha	
4	433.8	28.4	1434	6	E00631	E00631 DNA encodin	
5	432.2	28.3	1312	6	I02706	I02706 Sequence 1	
6	430.6	28.2	1584	9	BC011991	BC011991 Homo sapi	
7	429	28.1	1185	6	AR111410	AR111410 Sequence	
8	429	28.1	1352	6	AX335339	AX335339 Sequence	
9	429	28.1	1352	9	HUMALATM	K01396 Human alpha	
10	429	28.1	1371	6	AX335338	AX335338 Sequence	
11	429	28.1	1371	9	HSATPR1	X01683 Human mRNA	
12	429	28.1	1399	9	AK026174	AK026174 Homo sapi	
13	429	28.1	1434	6	E00195	E00195 cDNA encodi	
14	429	28.1	1434	6	I04196	I04196 Sequence 3	
15	429	28.1	1434	6	I04272	I04272 Sequence 1	
16	429	28.1	1434	6	I07849	I07849 Sequence 2	
17	429	28.1	1435	6	AX019486	AX019486 Sequence	
18	429	28.1	2478	17	AF130068	AF130068 Homo sapi	
19	427.4	28.0	1337	9	HUMALAT2	J02619 Human 2 typ	
20	427.4	28.0	1378	6	I02398	I02398 Sequence 8	
21	427.4	28.0	1431	9	BC015642	BC015642 Homo sapi	
22	427.4	28.0	2571	17	AF113676	AF113676 Homo sapi	
23	425.8	27.9	1299	6	I07949	I07949 Sequence 2	
24	425.8	27.9	1308	9	BABAIATA	J00321 Papio hamad	
25	425.8	27.9	1317	6	I00556	I00556 Sequence 2	
26	425.8	27.9	1378	6	I03509	I03509 Sequence 6	
27	425.8	27.9	1378	6	I07947	I07947 Sequence 4	
28	425.8	27.9	1434	6	I01352	I01352 Sequence 1	
29	425.4	27.9	1185	6	A01846	A01846 Human mRNA	
30	424.2	27.8	1378	6	I01227	I01227 Sequence 2	
31	417	27.3	1191	9	AB004044	AB004044 Cercopith	
32	407	26.7	1390	6	AX202089	AX202089 Sequence	
33	404.6	26.5	1356	6	I36163	I36163 Sequence 16	
34	403	26.4	1356	6	I36161	I36161 Sequence 12	
35	402.6	26.4	2013	6	AX472008	AX472008 Sequence 18	
36	399.8	26.2	1356	6	I36164	I36164 Sequence 18	
37	390.2	25.6	1351	10	AB000550	AB000550 Spermophil	
38	387	25.4	1372	10	AB000552	AB000552 Calloschi	
39	385.4	25.3	1351	4	SSANTIELA	X88780 S.scrofa mR	
40	378	24.8	1306	10	RATATRA1	M32247 rat alpha-1	
41	377.2	24.7	1352	10	S77822	S77822 alpha-1-ant	
42	376.4	24.7	1380	10	RATAPI	D06675 Rat mRNA fo	
43	372.8	24.4	1343	10	AB000546	AB000546 Tamias si	
44	372.6	24.4	1242	6	E13268	E13268 Temias asia	
45	370.6	24.3	1380	4	BTAIAT	X63129 B.taurus mR	

Pred. No. is the number of results predicted by chance to have a

QY	971	TGTTACTGAAGAAGCTCCATTAAAAATTGAGTAAAGCTGTTTCACAAAGCCGCTCTTAAGTAT	1030
Db	1076	CGTGACGAGGAGGCCCCCTGAAGCTCTCCAAGCGGTGCAAAAGCGGTGCTCAGAT	1135
QY	1031	TGATGAAAGGGTACCAGGCGCGCGCTATGTTCTCTGGAAGCTATTCGAATGAGCAT	1090
Db	1136	CGACGAGAAGGGACGGAAGTGCCTGGGGCCATGTTCTTGAGGCCATCCCATGTCAT	1195
QY	1091	TCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTCTTGATGATCGAGCAGACATAA	1150
Db	1196	CCGCGCCGAGGTCAGATTCACACGCCCTTCGTTCTCTGATGATCGAGCAGACAGAA	1255
QY	1151	AAGCCCATTTGTTATGGTAAAGTTGTCAACCCCAACTCAGAAG	1193
Db	1256	GAGCCCTCTTCATGGGGAAGTGTCAACCCCAAGGAGAAG	1298
RESULT 2			
AR111411		1185 bp	DNA
LOCUS		Sequence 3 from patent US 6127145.	linear
DEFINITION		AR111411	
ACCESSION		AR111411	
VERSION		AR111411.1	GI:12828259
KEYWORDS		Unknown.	
SOURCE		Unknown.	
ORGANISM		Unclassified.	
REFERENCE		1 (bases 1 to 1185)	
AUTHORS		Sutliff, T.D. and Rodriguez, R.L.	
TITLE		Production of -alpha- sub.1 -antitrypsin in plants	
JOURNAL		Patent: US 6127145-A 3 03-OCT-2000;	
FEATURES		Location/Qualifiers	
		1..1185	
BASE COUNT		276 a 396 c 334 g 179 t	
ORIGIN			
Query Match		41.2%;	Score 628.4; DB 6; Length 1185;
Best Local Similarity		70.7%;	Pred. No. 7.9e-124;
Matches 836;		Conservative 0;	Mismatches 346; Indels 0; Gaps 0;
QY	12	GAGAGCCCTCAAGGCGACGCGCTCAAAAAACCGACACACAGTCATCAGACCCAGACCAT	71
Db	1	GAGAGCCGCGAGGCGACGCGCCGACAGAGACCGACACGACCCAGCAGCAGGACAC	60
QY	72	CCGACTTTTAATAAATTACTCCAAATTTAGCGAATTTGCTTTTCTTTGTATAGACAA	131
Db	61	CCGAGTTCAACAGATTCACCCGGAATTTGGCGAATTCGCTCTCAGCTGTACCGCCAG	120
QY	132	TTAGCTCATCAAAAGTAATCTACTAAATTTTTTTTGTCTGTTCTTTATGCCACTGCT	191
Db	121	CTCGGACACAGTCCAACTCCACCAATCTTCTTCAGCCCGGTGAGCATCCGCCACGCC	180
QY	192	TTCGCCATGTGTTTAGTACTAAAGCCGATACCCATGACGAGATTTTGAAGGTTTA	251
Db	181	TTGCCATGTCTGCTTGGGTACCAAGGGGACCCACGACGAGATCTCCAGGGGTG	240
QY	252	AACTTTAAATTTGACGGAATCCCAAGAGCCCAATTTACAGAGGTTTTCAGAGTTGTTG	311
Db	241	AACTTCAACTGACGAGATCCCGGAGGCGCAGATCCACGAGGCTTCAGGAGTGTCTC	300
QY	312	AGAACTTTGAATCACTGATTTCTCAATTCGAATTAACCTGTTAACTGTTAACTGTTT	371
Db	301	AGGAGCTCAACACGCGGACTCCAGCTCCAGCTCACCACCGGACCGGGCTCTTCCTG	360
QY	372	TCTAAGGTTTAAATTTGGTTGACAAATTCCTAGAAGACTCAAGAACTATATCATAGT	431
Db	361	TCCGAGGCGCTCAAGCTCGTCGATAAGTTCTTGAGGAGCTGAAGAAGCTTACCACCTC	420
QY	432	GAGGCTTTTACCGTTAAATTTTGGTGATCTGAGAGACTTAAAGCAATTAATGATTAT	491
Db	421	GAGGCTTACCGTTCAACTTCGGGGACACGAGGAGGACGAGGAGGAGGAGGACTCAAC	480

Db	504	GAAGCCCTTCACCTGTCACAACTTCGGGATCACGAAGAGGCCAAGAAACAGATCAACGATTAC	563
Qy	492	GTGTGAAAGGACCCAGGGTAAGATCGTTGCACCTAGTTAAAGAAATTAGATCTGATACC	551
Db	564	GTGGAGAGGGTACTCAGGGGAAATTTGTGATTTGGTCAGAGAGCTTGACAGAGACACA	623
Qy	552	GTCTTCGCACCTAGTTAACTATATTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT	611
Db	624	GTTTTTGCTCTGCTGAATTACATCTCTTTTAAAGGCAAAATGGGAGAGACCTTTTGAAGTC	683
Qy	612	AAAGATACTGAAGAGAGATTTTCATGCTTCATCAAGTTACTACTGTCAAAGTTCACAAATG	671
Db	684	AAGGACACCGAGGAGGAGACTTCCACGTGGACCAAGGTGACCCCGTGAAGGTCCCTATG	743
Qy	672	ATGAAAGAGACTGGGTATGTTCAATATTCAACATTCGAAAAATTAAGTTCTTGGGTCTTA	731
Db	744	ATCAAGCGTTTAGGCATGTTTAACTCCAGCACTGTAAGAAGCTGTCCACGCTGGGTACTG	803
Qy	732	TTAATGAAGTATTTAGGTAAAGCTACTGCTATTTTTTTTACCAGACAGAGTAAAGCTT	791
Db	804	CTAATGAATACCTGGGCAATGCCACGCCCATCTCTTCTTACCTGATGAGGGGAAACTA	863
Qy	792	CAACATTTAGAGAAATGAGTTGACTCATGACATTTATTCTAAATTTTTTAGAGAACGAGGAT	851
Db	864	CAGCACTCGAAAAATGAACCTCACCCACGATATCATCAACCAAGTCTCTGGAATGAAGAC	923
Qy	852	CGTCTAGAGCGTTCCTGCACTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA	911
Db	924	AGAAGTCTCGCCAGCTTACATTTACCCAAAGCTGCCAATCTAGTGAACCTATGATCTGAAG	983
Qy	912	TCGTGTTTTAGGCGAGTTAGGTATTACCAAAGTTTTTTTCTAACGGTGCAGATTTGAGTGGT	971
Db	984	AGCGTCTGGTCACTGGCGCATCACTAAGGCTCTTCACCAATGGGCTGACCTCTCGGG	1043
Qy	972	GTTACTGAGAGAGCTCCATTAATTTAGTAAAGCTGTTTCAAAAGCGGCTCTTAACTATT	1031
Db	1044	GTCAAGAGAGGAGCACCCTGAAAGCTCTCAAGGCGGTGCATAGGCTGTGTGACCATC	1103
Qy	1032	GATGAAAAGGTTACCGAGGCGCGCGCTATGTTCTCTGGAAGCTATTCCAATGAGCATT	1091
Db	1104	GACGAGAGGGNACTGAAGCTGCTGGGCCATGTTTTTAGAGGCCATACCAATGCTCTATC	1163
Qy	1092	CCACGAGAGTTAAATTTAATAACCATTCGTTTTTCTGATGATCGAGCAGACACTAAA	1151
Db	1164	CCCCGAGAGTCAAGTTCACAAACCCCTTGCTCTTTAATGATGACAAAAATACCAAG	1223
Qy	1152	AGCCCATGTTTATGGTAAAGTTGTCACACCAACTCAGAA	1192
Db	1224	TCCTCCCTCTTATGGGAAAAGTGGTGAATCCCAACCCAAA	1264
RESULT 4	E00631	E00631	
LOCUS		1434 bp	RNA linear
DEFINITION		human alpha-1 antitrypsin	mutain.
ACCESSION		E00631	
VERSION		E00631.1	GI:2168910
KEYWORDS		JP 1986012289-A/1.	
SOURCE		Homo sapiens.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostom	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE		1 (bases 1 to 1434)	
JOURNAL		Maadaretsuto, W.I. and Guren, H.K.	
COMMENT		REGION SPECIFIC MUTATION INDUCTION IN ALPHA-1-ANTI-TRYPSIN	
		Patent: JP 1986012289-A 1 20-JAN-1986;	
		CHIMO JIENETEITSUKUSU INC	
		OS human	
		PN JP 1986012289-A/1	
		PD 20-JAN-1986	
		PF 14-MAR-1985 JP 1985051553	
		PR 14-MAR-1984 US 84 589410, 07-MAR-1985 US 85 709382	

QY 1092 CCACGAGAGTAAATTAATAACCAATTCGTTTTCTGTGATCGAGCAGACACTAA 1151
 Db 1180 CGCCCGAGGTCAAGTTCACAAACCCCTTGTCTCTTAATGATGAACAAATACCAAG 1239
 QY 1152 AGCCCATGTTTATGGTGAAGTTGTCAACCAACTCAGAA 1192
 Db 1240 TCTCCCTCTTCTGATGGAAGAGTGTGAATCCCAACCAAAA 1280

RESULT 6
 BC011991 1584 bp mRNA linear PRI 02-AUG-2001
 LOCUS Homo sapiens, Similar to serine (or cysteine) proteinase inhibitor,
 DEFINITION clade A (alpha-1 antiproteinase, antitrypsin), member 1, clone
 MGC-9222 IMAGE:3859644, mRNA, complete cds.
 ACCESSION BC011991
 VERSION BC011991.1 GI:15080498
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1584)
 Strausberg, R.
 Direct Submission
 Submitted (30-JUL-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalon@bcm.tmc.edu
 Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 21 Row: d Column: 6
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6855600.
 Location/Qualifiers
 1..1584
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:9222 IMAGE:3859644"
 /tissue_type="Ovary, adenocarcinoma"
 /clone_lib="NIH_MGC_66"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 233..1489
 /codon_start=1
 /product="Similar to serine (or cysteine) proteinase
 inhibitor, clade A (alpha-1 antiproteinase, antitrypsin),
 member 1"
 /db_xref="GI:15080498"
 /translation="MPSVSVSGILLLAGLCLVPVSLAEDPQGDAAQKDTFSHRDQDH
 PTFNKITNIAFEAFSLYROLAHSNSTNIFSPVSTATAFAMLSLGTAKDTHDEILE
 GUNFLNTEIPAQIHEGQFELLRLNQPSQLQLTGTNGLELSGLKVDKFLDVKK
 LYHSAFTVTEEDFEAKKQINDYVEKGTQGIQVDLVKELDRDVFALVNYLFEKWK
 ERPFVKQTEEDFEDFQVTVKPMKRLGMFIQVCKLSSWVLLMNYLGNATAF
 FLPDGKLSQHNELTHDITKFLNEDRRSASLHLPKLSITGYDLKSLVGLGKITK
 VPSNCAQLSGVTEAPLKLKSAVKHVLITIDEKGTAAAGAMFLEAIPMSIPPEVKFNK
 PFVFLMIQNTKSPFLFMGKVNPTQK"

FEATURES
 source

CDS

BASE COUNT 410 a 459 c 379 g 336 t
 ORIGIN
 Query Match 28.2%; Score 430.6; DB 9; Length 1584;
 Best Local Similarity 60.3%; Pred. No. 1.2e-81;
 Matches 712; Conservative 0; Mismatches 469; Indels 0; Gaps 0;
 QY 12 GAAGACCCCTCAAGGGACGCCCTCAAAAACCGACACAGTCATCAGCAACACCAT 71
 Db 305 GAGGATCCCGGAGATGCTGCCAGAAGACAGATACATCCCAACATGATCAGATCAC 364
 QY 72 CGACATTTTAAATAAATTAATCTCCAAATTTAGCCGAATTTGCTTTTCTTTGTATAGCAA 131
 Db 365 CCACCTTCAACAGATACCCCAACCTTGGCTGAGTTCGCCCTTACGCTATACCGCCAG 424
 QY 132 TTAGCTCATCAAGAGTAATTTCTATCAATTTTATTTAGTCTCTGTTTCTATTGCCACTGCT 191
 Db 425 CTGGCACACCAAGTCCCAACAGCAACCAATATCTTCTTCCGCCAGTGAGCATCGCTACAGCC 484
 QY 192 TTCGCCATGTTGATTTAGGTACTAAAGCCGATACCCATGACGAGATTTTGAAGGTTTA 251
 Db 485 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACATCCATGAAGGCTTCCAGGAACCTCCCTC 544
 QY 252 AACTTTAAATTTGACCGAAATCCCAAGAGCCCAAAATTCACGAGGGTTTCAAGAGTTGTTG 311
 Db 545 AATTTCAACCTCAGGAGATTCGGGAGGCTCAGATCCATGAAGGCTTCCAGGAACCTCCCTC 604
 QY 312 AGAATTTGAATCAACCTGATTTCTCAATTTGCAATTAACCTACTGTTAAACGTTTATTTTGG 371
 Db 605 CGTACCTCTCAACCAAGCCAGACAGCCAGCTCCAGCTCACCACCGGCAATGGCTTGTTCCTC 564
 QY 372 TCTGAAGGTTTAAATTTGGTTGACAAATTCCTAGAAGAGCTCAAGAAACTATATATAGT 431
 Db 665 AGCAGGGCTCAGCTAGTGGATAGTTTTCGAGGATGTTTAAAGAGTTGTACCACTCA 724
 QY 432 GAGGCTTTTACCCTTAATTTTGGTGATCTGAGGAGCTTAAAGCAAAATTAATGATTA 491
 Db 725 GAAGCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTA 784
 QY 492 GTTGAGAAAGGCCACCCAGGTAAGATCGTTGACCTAGTTTAAAGCAATTAAGATCGTGATAC 551
 Db 785 GTGAGAAAGGCTACTCAAGGGAAATTTGGTGAATTTGGTCAAGGAGCTTGACAGAGACACA 844
 QY 552 GTCTTCCACTAGTTAACTATATTTTCAAGGGTAAGTGGGAACGTCTCTTCGAGGTT 611
 Db 845 GTTTTCTGCTGTGATTAATCATCTTCTTAAAGCAAAATGGGAGAGACCTCTTGAAGTC 904
 QY 612 AAAGATACCTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAGTTCCAATG 671
 Db 905 AAGGACACCGGAGGAAGAGGACTTCCACGTGGACCAAGGTGACCCCGGTGAAGGTGCCTATG 964
 QY 672 ATGAAAGACTGGGTATGTTCAATATTCAACATTTGCAAAAAATTAAGTTCTTGGGTCTTA 731
 Db 965 ATGAAGCTTTTAGCATGTTTAAACATCCAGCTGTGAAGAGCTGCCACCTGGGTCG 1024
 QY 732 TTAATGAAGTATTTAGTAAAGCTTACTGCTATTTTATTTTACCAGACAGAGTAAAGCTT 791
 Db 1025 CTGATGAATACCTGGGCAATGCCACCGCATCTTCTTCTCTGCTGATGAGGGGAAACTA 1084
 QY 792 CAACATTTAGAGAATAGTTGACTCATGACATTAATTAATTAATTTTATAGAACAGGAT 851
 Db 1085 CAGCACCTGGAAAAATGAATCACCACAGATATCATCAAGTTCTCTGGAAAAATGAAGAC 1144
 QY 852 CGTGTGACGCTTCTCTGACCTCGCAAGTTAAGTATACCGGTACTTACGACTTAAAA 911
 Db 1145 AGAAGGCTCGCAGCTTACATTTACCAAACTGTCCATTAAGTGAACCTATGATCTGAAG 1204
 QY 912 TCTGTTTTAGGCCAGTTAGGTATTTACCAAAAGTTTTTCTTAACGGTGCCGATTTGAGTGGT 971
 Db 1205 AGCGTCTGGTCAACTGGGCATCCTAAGGTCTTCTCAGCAATGGGCGCTGACCTCTCCGGG 1264
 QY 972 GTTACTGAAGAGCTCCATTTAAATTTGAGTAAAGCTGTTCAAAAGCCGCTTAACTATT 1031

Best Local Similarity 60.2%; Pred. No. 2.6e-81; Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

QY	12	GAAGACCTCAAGCGAGCGCGCTCAAAACCGACACAGTCAATCACGACCAAGACCAT	71
Db	92	GAGGATCCCGAGGAGATGCTGCCAGACAGATACATCCACCATGATCAGGATCAC	151
QY	72	CCGACTTTAATAAAATTAATCTCAAAATTTAGCCGAATTTGCTTTTCTTGTATAGACAA	131
Db	152	CCACCTTCACACAGATCACCCCAACCTGGCTGAGTTCGCCCTATACCCGCCAG	211
QY	132	TTAGCTCATCAAGTAATCTACTAATTTTTTTTGTAGTCTGTTTCTATGCCACTGCT	191
Db	212	CTGGCACACAGTCCACAGACCAATATCTTCTCCCGAGTGAGCATCGCTACAGCC	271
QY	192	TTCCGCTGTTGAGTTAGTACTTAAGCCGATACCCATCAGCAGATTTTGAAGAGTTTA	251
Db	272	TTTCAATGCTCTCCCTGGGGACCAAGCTGACACTCAGATGAATCTCGGAGGCGCTG	331
QY	252	AACCTTTAATTTGACGAAATCCAGAGCCCAATTCACGAGGGTTTTCAGAGTGTGTTG	311
Db	332	AATTTCAACCTCAGGAGATTCGGAGGCTCAGATCCATGAAGCTTCCAGGAACCTCTC	391
QY	312	AGAACTTTGAATCAACCTGATCTCAATTTGCAATTAATCTACTGTTAGCGGTTTATTTTG	371
Db	392	CGTACCCCTCAACACGACAGCAGCTCCAGCTGACCTACCCAGCAATGGCCTGTTCCCTC	451
QY	372	TCTGAAGTTTAAATGTTGACAAATTCCTAGAAGCTCAAGAACTATATCATAGT	431
Db	452	AGCGAGGCGCTGAAGCTAGTGGATGAATTTTGGAGGATGTTAAAGATGTGACCACTCA	511
QY	432	GAGGCTTTTACCGTTAATTTTGGTATGACTGAGGAAGCTAAAGCAAAATTAATGATAT	491
Db	512	GNAGCTTCACTGTCACTTCGGGACACCGAGGCAAGGCAAGAAACAGATCAAGGATAC	571
QY	492	GTGGAAGAGGCCACCGAGGTAAGATGTTGACCTAGTTAAAGAAATTAAGATCGTGATAC	551
Db	572	GTGGAGAAGGGTACTCAAGGGGAAATTTGGATTTTGGTCAAGGAGCTTGACAGAGACACA	631
QY	552	GTCCTCGCAGTATTAATATATTTTCAAGGTAAGTGGGACGTCCTTTTCGAGGTT	611
Db	632	GTTTTGTCTGTTGGAATTTACATCTCTTTTAAAGGCAATGGGAGAGACCTTTGAAGTC	691
QY	612	AAAGATCTGAAGAGGAAGATTTTCATGTTGATCAAGTACTACTGTCAAAAGTTCCAAATG	671
Db	692	AGGACACCGGAGAGAGACTTCCACGTGGACAGGTGACACCGTGAAGGTGCCTATG	751
QY	672	ATGAAAGACGTTGGTATGTTCAATATTCACATTTGCAAAATTAAGTCTTGGGCTTTA	731
Db	752	ATGAAGCGTTTAGGCATGTTTAAATCCAGCACTGTAAAGAGCTGTCCAGCTGGGTGCTG	811
QY	732	TTAATGAAGTATTTAGTAAAGCTACTCTATTTTTTTTTTACCAGACGAAGGTAAGCTT	791
Db	812	CTGATGAATACCTGGGCAATGCCACCGCATCTTCTTCCTGCTGATGAGGGGAACATA	871
QY	792	CAACATTTAGAGAATGAGTTGACTCATGACATTAATTAATTTTATAGAAACGAGGAT	851
Db	872	CAGCCTTGGAAATGAAGTCAACCAAGATATCATCAACAAAGTTCCTGGAAATGAAGAC	931
QY	852	CGTCGTAGCGCTTCTCTGACCTGCCAAAGTTAAGTATACCGGTACTTACGACTTAAAA	911
Db	932	AGAAGCTCTGCCAGCTTACATTTTACCCAACTGTCCATTTACTGGAACCTATGATCTGAAG	991
QY	912	TCTGTTTTAGGCGAGTTAGTATTAACAAATTTTTTCTTAACGGTGGCGATTTGAGTGGT	971
Db	992	AGCGCTCTGGGTCACTGGGCATCACTAGGCTCTTACAGATGGGCTGACCTCTCCGGG	1051
QY	972	GTTACTGAAGAGAGCTCCATTAATTAATTTAGTAAAGCTGTTCACAAAGCGCTCTTAACAT	1031
Db	1052	GTCACAGAGGAGCCCTGAGCTCTCAAGGCGGTGCATGAAGCTGTGCTGACCATC	1111
QY	1032	GATGAAAGGGTACGAGGCCCGCGGTATGTTCTTGGAAAGCTATTTCCATGAGCAT	1091

Db	1112	GACGAAAGGACTGAAGCTGCTGGGCCATGTTTTTAGAGGCCATACCATCTCTATC	1171
QY	1092	CCACAGAGTTAAATTTAATAAACCATTCGTTTCTGATGATCGAGCAGAACACTAAA	1151
Db	1172	CCCCCGAGGTCAAGTTCACAAACCCCTTGCTCTTCTTAATGATTTGAACAAATACCA	1231
QY	1152	AGCCCATTTGTTATGGTAAAGTTGTCACACCAACTCAGAA	1192
Db	1232	TCTCCCTCTTCATGGGAAAGTGTGAATCCACCCAAA	1272

RESULT 9
HUMALATM Human alpha-1-antitrypsin mRNA, complete cds.
LOCUS K01396.1 GI:177828
DEFINITION alpha-1-antitrypsin; antiprotease; antitrypsin.
ACCESSION Human liver, cDNA to mRNA, clones PULB1523 [1] and pTG603 [3].
VERSION
KEYWORDS
SOURCE Homo sapiens
ORGANISM
REFERENCE 1 (bases 1 to 1231)
AUTHORS Bollen, A., Herzog, A., Cravador, A., Herion, P., Chuchana, P., Vander Straten, A., Loriau, R., Jacobs, P. and van Elsen, A.
TITLE Cloning and expression in Escherichia coli of full-length complementary DNA coding for human alpha 1-antitrypsin
JOURNAL DNA 2 (4), 255-264 (1983)
MEDLINE 84107980
PUBMED 6319097
REFERENCE 2 (bases 1 to 1352)
AUTHORS Colau, B., Chuchana, P. and Bollen, A.
TITLE Revised sequence of full-length complementary DNA coding for human alpha 1-antitrypsin
JOURNAL DNA 3 (4), 327-330 (1984)
MEDLINE 85026667
PUBMED 633329
REFERENCE 3 (bases 95 to 286)
AUTHORS Courtney, M., Buchwalder, A., Tessier, L.H., Jaye, M., Benavente, A., Balland, A., Kohli, V., Lathé, R., Tolstoshev, P. and Lecocq, J.P.
TITLE High-level production of biologically active human alpha 1-antitrypsin in Escherichia coli
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 81 (3), 669-673 (1984)
MEDLINE 84144765
PUBMED 6322161
COMMENT Alpha-1-antitrypsin is synthesized in the liver and is a major constituent of plasma. It functions as an inhibitor of elastase, which degrades lung tissues. Hereditary deficiencies are fairly frequent, causing early lung degeneration and sometimes severe liver disorders. Defectives are particularly sensitive to pollution induced inflammation. Injection of anti-trypsin relieves the condition [1].

FEATURES	Location/Qualifiers
source	1. .1352 /organism="Homo sapiens" /db_xref="taxon:9606" /map="14q32.1"
gene	1. .1352 /gene="PI"
mRNA	<1. .1352 /gene="PI" /product="al-antitrypsin mRNA"
CDS	20. .1276 /gene="PI" /note="alpha-1-antitrypsin" /codon_start=1 /protein_id="AAB59375.1" /db_xref="GI:177828" /db_xref="GDB:G00-120-289" /translation="MPSSVSWGILLIAGLCCILPVSVLAEDPOGDAQAOKTDTSHDDQDH PTFKTTPNLAFAFSLYROLAQSNSNIFPSVTSIAFAFAMLSIGTKADHDELE GLNFNTEIPEAQIHEGFQELLRLNQDLSQLTGNGFLISEGLKLVDRKLEVDYK LYHSEFTVNFQDTEAKQINDYVEKGTQGIKVIDLVKELDRDTFVFLVNYFFKQK

RRPREVKDEDEHVDVTVKVPMMKRLGFMENIOHCKKLSSNVLLMKYLGNATAIF
FLPDGKQLHLENELTHDIIITKFLNEDRRSASLHPLKLSITGYIDLKSVLQUGITK
VFSNGADUSGVTEAPLKLKSLRAVHKAVLTIDEKTEAGAMFLEAIPMSIPPEVKFNK
PFVFLMIDONTKSPFLFMKGKVVNPTQK*
20. .91

sig_peptide

/gene="PI"
/note="alpha-1-antitrypsin signal peptide"
92. 1273
/gene="PI"
/product="alpha-1-antitrypsin mature peptide"
241
/gene="PI"
/citation=[1]
286
/gene="PI"
/citation=[3]
/replaces=""
435. .439
/gene="PI"
/citation=[1]
748
/gene="PI"
/citation=[1]
784
/gene="PI"
/citation=[1]
787
/gene="PI"
/citation=[1]
799. .802
/gene="PI"
/citation=[1]
837
/gene="PI"
/citation=[1]
1114

conflict

old_sequence

old_sequence

old_sequence

old_sequence

old_sequence

old_sequence

old_sequence

BASE COUNT 349 a 386 c 325 g 292 t
ORIGIN 90 bp upstream of Ddel site.

Query Match 28.1%; Score 429; DB 9; Length 1352;

Best Local Similarity 60.2%; Pred. No. 2.6e-81;

Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

QY 12 GAAGACCTCAAGGCGACCGCTCAAAAACCGACACAGTCATCAGCAACAGACCAT 71

DB 92 GAGGATCCCGGAGAGTGCTGCCAGAGACAGATACATCCACCAGTACAGATCAC 151

QY 72 CCGACTTTTAAATAAATACTCCAATTTAGCGGAATTTGCTTTTCTTTGTATAGACAA 131

DB 152 CCACCTTCAACAAGATCACCCCAACCTGGCTGAGTTGCGCTTACCGCAG 211

QY 132 TTAGTCATCAAGTAATTTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191

DB 212 CTGGCACACCTCCCAACAGCAACCAATATCTTCTTCTCCCGAGTGAGCATGCTACAGCC 271

QY 192 TTCGCCATGTCAGTTTAGTACTAAAGCCGATCCCATGACGAGATTTTGAAGGTTTA 251

DB 272 TTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCAGATGAATCTCGAGGCGCTG 331

QY 252 AACTTTAATTTTACCGAATCCAGAGCCCAAAATTCAGAGGTTTTCAGAGTTGTTG 311

DB 332 AATTTCAACCTCAGGAGATCCCGAGGCTCAGATCCCATGAAGGTTCCAGAACTCCTC 391

QY 312 AGAAGCTTTGAATCAACCTGATTTCTCAATTTGAATTTTAACTACTGTTAGCGGTTTATTG 371

DB 392 CGTACCTCAACAGCCAGACAGCCAGCTCCAGCTGACCCAGGCAATGGCTGTCTC 451

QY 372 TCTGAAGGTTTAAATTTGTTGACAAATTCCTAGAGAGCTGCAAGAACTATATAGT 431

DB 452 AGCGAGGCGCTGAAGCTAGTGAATTTTGGAGGATGTTTAAAGATTTGTACCACTCA 511

QY 432 GAGGCTTTTACCCTTAATTTTGGTGATCTGAGGAGAGCTTAAAGCAAAATTAATGATTAT 491
DB 512 GAAGCCTTCTACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 571
QY 492 GTTGAGAAAGCCACCCAGGTAAAGATCGTTGACCTAGTTAAAGAAATTAGATCGTGATACC 551
DB 572 GTGAGAAAGGTACTCAAGGAAATTTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 631
QY 552 GTCCTCCACTAGTTAACTATATTTTTTCAAGGTAAGTGGGAACGCTCTTTTCGAGGTT 611
DB 632 GTTTTTGCTGCTGGTGAATTTACATCTTCTTAAAGCAAAATGGGAGAGACCTTTTGAAGTC 691
QY 612 AAAGATCTGAAGAGGAAGATTTTCATGTGTGATCAAGTTACTACTGTCAAGTTCCAAATG 671
DB 692 AAGGACACCGAGGAGAGGAGCTTCCAGCTGGACGAGTGACCCAGTGAAGGTGCGCTATG 751
QY 672 ATGAAAGAGCTGGGTATGTTCAATATTTCAACATTTCAACATTTGCAAAAAATTTAAGTTCTTGGGTCTTA 731
DB 752 ATGAAGCTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTCTG 811
QY 732 TTAATGAAGTATTTAGGTAAAGCTGCTGCTATTTTTTTTTTACAGAGCAAGGTAAAGCTT 791
DB 812 CTGATGAATACCTGGGCAATGCCACGCCATCTTCTTCCCTGCTGATGAGGGGAAACTA 871
QY 792 CAACATTTAGAGAAATGAGTTGACATCATGACATTAATTAATTTTGTAGAGAACGAGGAT 851
DB 872 CAGCAGCTGGAAATGAATCAACCCAGATATCATCAAGTTTCTTGGAAAAATGAAGAC 931
QY 852 CGTGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCATCCGCTACTTACGACTTAAAA 911
DB 932 AGAAGCTGCGCAGCTTACATTTACCCCAACTGTCCATTTACTGGAACCTATGATCTGAAG 991
QY 912 TCTGTTTTAGGCCAGTTAGGTATTTACCAAGTTTTTTTTTCTAACGGTGCCTGAGTGGT 971
DB 992 AGCGTCTGGTCAACTGGGCATCACTAAGTCTTTTACAGCAATGGGGTGACCTCTCGGG 1051
QY 972 GTTACTGAAGAGCTCCATTAATAATTCAGTAAAGCTCTTCACAAAGCCGCTTAACTATT 1031
DB 1052 GTCAGAGAGGAGCCCGCTGAGCTCTCCAGGCCGCTGATAGGCTGTGCTGACCATC 1111
QY 1032 GATGAAAGGTTACCGAGCGCGCGGCTATGTTTCTTGGAAAGCTATTTCAATGAGCAT 1091
DB 1112 GACGAGAAAGGAGCTGAAGCTGCTGGGCCATGTTTTTAGAGGCCATACCATGCTCTATC 1171
QY 1092 CCACCAAGAGTTAAATTTAATAAACCAATTCGTTTTCTGTATGATCGAGGAGAACACTAAA 1151
DB 1172 CCCCCGAGGTCAGATTTCAACAAACCTTTTGTCTTTAATGATTGAACAAAAATACCAAG 1231
QY 1152 AGCCCATTTGTTATGGGTAAGTTGTCAACCCCAACTCAGAA 1192
DB 1232 TCTCCCTCTTCATGGGAAAAAGTGGTGAATCCCAACCCAAAA 1272

RESULT 10

AX335338

LOCUS

DEFINITION

AX335338

ACCESSION

AX335338.1

VERSION

GI:18126057

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

AX335338 Sequence 5847 from Patent WO0194629.
AX335338
AX335338.1 GI:18126057

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D. R. and Weaver, Z.

Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 5847 13-DEC-2001;
Avalon Pharmaceuticals (US)

Location/Qualifiers
1. .1371

```
Query Match 28.1%; Score 429; DB 6; Length 1371;
Best Local Similarity 60.2%; Pred. No. 2.6e-81;
Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

BASE COUNT 350 a 388 c 340 g 293 t
ORIGIN

QY 12 GAAGACCTCAAGCGGAGCGCTCAAAAACGACACACAGTCATCAGCAGCAGCAGCAT 71
DB 111 GAGGATCCCCAGGAGATGCTGCGCAGAGACATATATCCACCACATGATCAGGATCAC 170
QY 72 CCGACTTTTAATAAAATTAATCTCCAAATTTAGCCGAATTTGCTTTTCTTTGTTATAGACAA 131
DB 171 CCAACCTTCAACAGATCAACCCCACTGGCTGAGTTCGCTTACGCTATACCGCCAG 230
QY 132 TTAGCTCATCAAGTAATCTTACTAAGATTTTTTTTGTCTCTGTTCTATGCGACTGCT 191
DB 231 CTGGCACACCACTCAACAGCAGCAATATCTTCTCCCGAGTGAGCATCGCTACAGCC 290
QY 192 TTCGCCATGTTAGTTAGTAAAGCGATACCCATGACGAGATTTTAGAAGGTTTA 251
DB 291 TTGCAATGCTCTCCCTGGGACCAAGCTGACACTCAGATGAATCCTTGGAGGCGCTG 350
QY 252 AACTTTAATTTGACCGAAATCCCAAGACCCCAATTCACGAGGGTTTTCAGAGATTTG 311
DB 351 AATTTCAACCTCAGGAGATTCGGAGGCTCAGATCCATCAAGGCTTCCAGGAACCTCCTC 410
QY 312 AGAAGCTTGAATCAACCTGATCTCAATGCAATTAAGTCTGTAACGGTTATTTTG 371
DB 411 CGTACCTTCAACCAAGCAGACGAGCTCCAGCTGACCGCAATGGCTGTTCCCTC 470
QY 372 TCTGAAGTTTAAATGTTGACAAATTCCTAGAAGCTCAAGAACTATATCATAGT 431
DB 471 AGCGAGGCTTGAAGTGTGATGATGATTTTGGAGATGTTAAAGATGTTACCACTCA 530
QY 432 GAGGCTTTTACCGTTAATTTTGGTGATCTAGTGAAGCTTAAAGACAAATTAATGATTAT 491
DB 531 GAAGCTTCACTGTCACTTCCGGGACACCGGAGGCTCAAGAGGCAAGAAACAGATCAAGATATC 590
QY 492 GTTGAGAAGCAGCCAGGCTTAAGATGTTGACCTAGTTAAAGATTAAGATGATGATACC 551
DB 591 GTGGAGAAGGCTACTCAAGGAAATTTGTGATTTGGTCAAGGAGCTTGACAGAGACACA 650
QY 552 GTCTTCGCACGTAGTAACTATATTTTTCAGGGTAAAGTGGGACGCTCTTCGAGGTT 611
DB 651 GTTTTGTCTGTTGAATTAATCTCTTTAAAGGCAATGGGAGAGACCTTTTGAAGTC 710
QY 612 AAAGATCTGAAGAGAGATTTTCATGTTGATCAAGTTACTACTGTCAAGATTCGAATG 671
DB 711 AAGGACACCGAGGAGGAGCTTCCAGCTGGACGAGGTGACCCGTTGAAGGTGCCATG 770
QY 672 ATGAAAAGACTGGGTATGTTCAATATTCACATTTGCAAAATTAAGTTCTTGGGTCTTA 731
DB 771 ATGAGCGCTTAGCACTGTTTAACATCCAGCACTGTAGAAGCTGTCCAGCTGGGTGCTG 830
QY 732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGAGGATGAGCTT 791
DB 831 CTGATGAATACCTGGGCAATGCCACGCCCATCTTCTCTCCCTGATGAGGGGAAACTA 890
QY 792 CACATTTAGAGATGAGCTTCACTCATGACATTTACTAATTTTATAGAACGAGGAT 851
DB 891 CAGCACCCTGGAAATGAAGTCAACCCAGATATCATCAAGATTCCTGGAAATGAAGAC 950
QY 852 CGTCTAGCGGTTCTCTGACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAA 911
DB 951 AGAAGGCTGCCAGCTTACATTTACCCAACTGCTCATTTACTGGAACCTATGATCTGAAG 1010
QY 912 TCTGTTTATAGCCAGGTTAGGTAATACCAAGTTTTTTCTTAACCGTGCGGATTTGAGTGGT 971
DB 1011 AGCGTCTGGGTCAACTGGGATCACTAAGGCTTCTAGCAATGGGGTGTGACCTCTCCGGG 1070

/organism="Homo sapiens"
/db_xref="taxon:9606"
293 t
340 g
388 c
350 a
```


Query Match		28.1%;	Score 429;	DB 9;	Length 1399;
Best Local Similarity		60.2%;	Pred. No. 2.6e-81;		
Matches 711;		Conservative	0;	Mismatches 470;	Indels 0; Gaps 0;
QY	12	GAAGACCTCTAAGCGGAGCGCTCAAAAACGACACACAGTATCATCAGCAGCAAGACCAT	71		
Db	116	GAGGATCCCCAGGAGATGCTGCCAGAAAGACAGATACATCCACCACATGATCAGGATCAC	175		
QY	72	CCGACTTTTAATAAATTTACTCCAAATTTAGCCGAATTTGCTTTTCTTTGTATAGACAA	131		
Db	176	CCAACTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCTTCAGCCTATATCTGCCAG	235		
QY	132	TTAGCTCATCAAGTAATTTCTACTAAACATTTTTTTTAGTCTCTGTTTCTTATGCCACTGCT	191		
Db	236	CTGGCACACCACTCCACAGCACCAATATCTCTCTCCCGCAGTGACATCGCTACAGCC	295		
QY	192	TTCCGCTGTTGAGTTTAGGTACTTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA	251		
Db	296	TTTGC AATGCTCTCCCTGGGGACCAAGGCTGACACTCACATGAATCCTGGAGGGGCTG	355		
QY	252	AACTTTAATTTGACCGAATCCCAAGACCCAAATTCACGAGGGTTTTCAGAGATTGTTG	311		
Db	356	AATTTCAACCTCAGGAGATTCGGAGGCTCAGATCCATGAAGCTTCCAGGAACCTCCTC	415		
QY	312	AGAATTTTGAATCAACCTGATTCTCAATTCGAATTAACCTACTGTTAAGGTTTATTTTGG	371		
Db	416	CGTACCTCAACCAAGCCAGACGACGCTCCAGCTGACACCGGCAATGGCTGTTCCTC	475		
QY	372	TCTGAAGGTTTAAATTTGGTTGACAAATTCCTAGAAGACGTCAGAAACATATATCATAGT	431		
Db	476	AGGAGGGGCTGAAGCTAGTGATGAATGTTTTTGGAGGATGTTAAAGTTGTACCACTCA	535		
QY	432	GAGGCTTTTACCGTTAAATTTTGGTGATCTACTGAGGAAGCTAAAAAGCAAAATTAATGATTAT	491		
Db	536	GAAGCCTTCACTGTCAACTTCGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC	595		
QY	492	GTTGAGAAGGACCCAGGTAAGATCGTTGACCTAGTTTAAAGATTTAGATCGTGATACC	551		
Db	596	GTGGAGAGGGTACTCAAGSAAAAATTTGGATTGTCAGAGGAGCTTGACAGAGACACA	655		
QY	552	GTCTTCGCACGTAGTTAACTATATTTTTTCAAGGTAAGTGGGAAACGTCCTTTTCGAGGTT	611		
Db	656	GTTTTTGTCTGTTGAATTAATCTCTTTTAAAGCAATGGGAGACCCCTTTTGAAGTC	715		
QY	612	AAAGATCTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCCAATG	671		
Db	716	AAGGACACCGGAGGAGGACTTCCACGTGGACGAGGTGACCCCGTGAAGGTGCCTATG	775		
QY	672	ATGAAAAGACTGGGTATGTTCAATATTCAACATTCGCAAAATTAAGTTCTTTGGGTCTTA	731		
Db	776	ATGAAGCGTTTAGGCATGTTTAACTCCAGCACTGTAGAAGCTGTCCAGCTGGGTGCTG	835		
QY	732	TTAATGAAGTATTTAGGTAAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT	791		
Db	836	CTGATGAAATACCTGGGCAATGCCACGCCATCTTCTTCTCCTGCTGATGAGGGGAAACTA	895		
QY	792	CAACATTTAGAGAATGAGTTGACTCATGACATTAATTAATAATTTTATAGAACGAGAT	851		
Db	896	CAGCACCTGGAAAATGAATCAACCCACGATATCATCAAAAGTTCCTGGAAAATGAAGAC	955		
QY	852	CGTCGTAGGCTTCTCTGCACCTGCCAAGTTAAGTATCACCGGTACTTACGACTTAAA	911		
Db	956	AGAAGGCTGCCAGCTTACATTTACCAAACTGTCCATTTACTGGAACCTATGATCTGAAG	1015		
QY	912	TCGTGTTTATAGCCAGTTAGTATTAACAAAGTTTTTTTCTAACGGTGGCGGATTTGAGTGGT	971		
Db	1016	AGCGTCTCGGTTCAACTGGGATCACTAAGGTCTTTCAGCAATGGGGCTGACCTCTCCGGG	1075		
QY	972	GTTACTGAGAAAGCTCCATTAATAATTTAGTAAGCTGTTTCACAAAGCGGCTCTTAACTATT	1031		
Db	1076	GTCAAGAGGAGGACACCCCTGAAGCTCTCCAAAGGCGGTGCATGAAGGCTGTGCTGACCATC	1135		
QY	1032	GATGAAAAGGGTACCGAGGCCCGCGGCTATGTTTCTCTGGAAAGCTATTTCCAATGAGCAATT	1091		
Db	1136	GACGAGAAAGGACTGAAGCTGCTGGGCCATGTTTTTAGAGCCCATACCATGCTATC	1195		
QY	1092	CCACCAGAAAGTTAAATTTAATAAACCATTCGTTTCTGTGATGATCAGCAGAACACTAAA	1151		
Db	1196	CCCCCGAGGTCAAGTTCAACAAACCCCTTGTCTCTTAATGATTGAACAAAATACCAAG	1255		
QY	1152	AGCCCATGTTTATGGTGAAGTTGTCAACCCAACTCAGAA	1192		
Db	1256	TCTCCCTCTTTCATGGGAAAAGTGGTGAATCCCAACCAAAA	1296		
RESULT 13					
E00195					
LOCUS					
DEFINITION					
CDNA encoding human antitrypsin.					
ACCESSION					
E00195					
VERSION					
E00195.1 GI:2168491					
KEYWORDS					
JP 1984091886-A/1.					
SOURCE					
Homo sapiens.					
ORGANISM					
Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE					
1 (bases 1 to 1434)					
Guren,E.K. and Richiyaado,U.					
SUGAR LYSING PROMOTOR : PROTEASE INHIBITOR FOR DEVELOPMENT OF					
ADJUSTED PROTEIN					
Patent: JP 1984091886-A 1 26-MAY-1984;					
JIMOSU CORP					
COMMENT					
OS (human)					
PN JP 1984091886-A/1					
PD 26-MAY-1984					
PF 12-AUG-1983 JP 1983147907					
PR 13-AUG-1982 US 82 408099, 28-APR-1983 US 83 489406 PI					
GUREN EICHI KAWASAKI, RICHIIYAADO UTSUOBERII					
PC C12N15/00,A61K35/72,A61K37/64,C07G7/00,C07H21/04,C12N1/16, PC					
C12N9/99,					
PC C12P21/00,(C12N1/16,C12R1.865);					
CC strandedness: Double;					
CC topology: Linear;					
CC hypothetical: No;					
CC anti-sense: No;					
FH Key					
FH Location/Qualifiers					
FT 5'UTR					
FT CDS					
FT 3'UTR					
FT mRNA					
FT sig_peptide					
FT mat_peptide					
FT					
FEATURES					
source					
1. .1434					
/organism="Homo sapiens"					
/db_xref="taxon:9606"					
BASE COUNT					
369 a 432 c 340 g 293 t					
ORIGIN					
Query Match					
Best Local Similarity					
28.1%; Score 429; DB 6; Length 1434;					
Pred. No. 2.6e-81;					
Matches 711; Conservative					
0; Mismatches 470; Indels 0; Gaps 0;					
QY	12	GAAGACCTCTAAGCGGAGCGCTCAAAAACGACACACAGTATCATCAGCAGCAAGACCAT	71		
Db	119	GAGGATCCCCAGGAGATGCTGCCAGAAAGACAGATACATCCACCACATGATCAGGATCAC	178		
QY	72	CCGACTTTTAATAAATTTACTCCAAATTTAGCCGAATTTGCTTTTCTTTGTATAGACAA	131		
Db	179	CCAACTTCAACAAGATCACCCCCAACCTGGCTGAGTTGCGCTTCAGCCTATACCCCGAG	238		
QY	132	TTAGCTCATCAAAAGTAATTTCTACTAAACATTTTTTTTAGTCTCTTCTTATGCGCATGCT	191		
Db	239	CTGGCACACCACTCCACAGACACCAATATCTTCTTCCCGCAGTGAGCATCGCTACAGCC	298		


```
Db 839 CTGATGAATACCTGGGCAATGCCACCGCCATCTTCTTCTCCTGCTGATGAGGGAACTA 898
Qy 792 CAACATTTAGAGATAGTTGACTCATGACATTAATTAATAATTTTATAGAAACGAGGAT 851
Db 899 CAGCACCTGGAAATGAACCTCACCCACGATATCATACCAAGTTCCTGGAAATGAAGAC 958
Qy 852 CGTGTAGCGTCTCTGACCTCGCAAGATTAGTATACCCGGTACTTACGACTTTAAAA 911
Db 959 AGAAGGTCTGCAGCTTACATTTTACCCAACTGTCCATTTACTGGAACCTATGATCTGAAG 1018
Qy 912 TCTGTTTATAGCCAGTTAGTATTAACCAAGCTTTTCTTAACGGTGCCGATTTGAGTGGT 971
Db 1019 AGCGTCTGGGTCACTGGGCATCACTAAGTCTCTCAGCAATGGGGCTGACCTCTCCGG 1078
Qy 972 GTTACTGAAGAGCTTCCATTAATAATTGATTAAGCTGTTCACAAAGCCGCTTTAACTATT 1031
Db 1079 GTACAGAGGAGGCCCTCAAGCTCTCCAAAGCCGTGCAATAAGGCTGTCTGACCATC 1138
Qy 1032 GATGAAAGGTTACCGAGCGCCGCGCTATGTTCTGGAGCTATTCAGAGCTATTCGAATG 1091
Db 1139 GACGAAAGGAGTGAAGCTGCTGGGGCCATGTTTATAGAGGCCATACCCATGCTATC 1198
Qy 1092 CCACAGAGTTAAATTTAATAAACCAATTCGTTTCTGATGATCGAGCAGAACACTAAA 1151
Db 1199 CCCCCGAGGTCAAGTTCACAAACCCCTTTGCTCTTCTTAATGATTGAACAATACCAAG 1258
Qy 1152 AGCCCATTTTATGGGTGAAGTTGTCAACCCAACTCAGAA 1192
Db 1259 TCTCCCTCTTCATGGGAAAGTGGTGAATCCACCCCAAAA 1299

RESULT 15
I04272
LOCUS I04272 1434 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent EP 0139383.
ACCESSION I04272
VERSION I04272.1 GI:591814
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1434)
AUTHORS Russell,P.R.
TITLE Method for expressing foreign genes in schizosaccharomyces pombe
and the use in therapeutic formulations of the products, DNA
constructs and transformant strains of schizosaccharomyces pombe
usable in such method and their preparation
Patent: EP 0139383-A1 1 02-MAY-1985;
JOURNAL Location/Qualifiers
FEATURES
source I..1434
BASE COUNT 369 a 432 c 340 g 293 t
ORIGIN

Query Match 28.1%; Score 429; DB 6; Length 1434;
Best Local Similarity 60.2%; Pred. No. 2.6e-81;
Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

Qy 12 GAAGACCTCAAGGCGAGCGCGCTCAAAAAACCGACACAGTCAATCAGCAACCAAGACCAT 71
Db 119 GAGGATCCCCAGGAGATGCTGCCCAAGACAGATACATCCACCATGATCAGGATCAC 178
Qy 72 CGGACTTTTAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 131
Db 179 CCAACCTTCAACAAGATCACCCCAACCTGGCTGAGTTTCGCTTCAGCCTATACGCCAG 238
Qy 132 TTAGCTCATCAAGTAATTTCTACTAATTTTCTTTAGTCTGTTCTTATTTGCCACTGCT 191
Db 239 CTGGCACACCACTCCAAACAGCAGCAACCAATATCTTCTTCCCGAGTGAGCATCGCTAC 298
Qy 192 TTTCGCATGTTGAGTTAGTACTAAAGCCGATACCCAGATGAGGATTTTAGAAGGTTTA 251
Db 299 TTTGCANTGCTCTCCTGGGGACCAAGGCTGACACTCAGATGAATCTCTGGAGGCGCTG 358
```

* . . *

```
Qy 252 AACTTTAAATTTGACCGAAATCCAGAGAGCCCAAAATTCACGAGGGTTTTCAAGAGTTGTG 311
Db 359 AATTTCAAACTCAGGAGATTCGGAGGCTCAGATTCATGAGGCTTCAGGAACCTCTC 418
Qy 312 AGAACCTTTGAATCAACCTGATTTCTCAATTTGCAATTAACCTACTGTTAAACGGTTTATTT 371
Db 419 CGTACCTCTCAACAGCAGACAGCCAGCTCCAGCTGACCACCGCAATGGCGTGTCTCTC 478
Qy 372 TCTGAAGGTTTAAATTTGTTGACAAATTCCTTAGAAGAGCTCAAGAAATATATATATAGT 431
Db 479 AGCGAGGCGCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTA 538
Qy 432 GAGGCTTTTACCGTTAATTTTGGTGATCTGAGGAAGCTTAAAGCAAAATTAATGATTAT 491
Db 539 GAAGCCTCTCAGCTGCACTTCGAGGACACCGAGAGGCCCAAGAAACAGATCAACGATAC 598
Qy 492 GTTGAAGAAAGGCCACCCAGGCTAGATCGTTGACCTAGTTTAAAGAAATTAGATCGTGATCC 551
Db 599 GTGGAGAGGGTACTCAAGGAAATTTGGGATTTTGGTCAAGGAGCTTCACAGAGACACA 658
Qy 552 GTCTTCGCACTAGTTAACTATATTTTTCAGGGTAAAGTGGGAGAGCTCTTTCAGAGTT 611
Db 659 GTTTTGTCTCTGTGTAATACATCTTTTAAAGGCAAAATGGGAGAGACCCCTTTGAAGTC 718
Qy 612 AAGTACTCAAGAGAGAGATTTTCATGTTGATCAAGTTTACTACTCTCAAAGTTTCCAATG 671
Db 719 AAGGACACCGAGGAGAGGACTTCCACGTGGACCGAGGTGACCACCGTGAAGGTGCGTATG 778
Qy 672 ATGAAAGAGCTGGGTATGTTTCAATATTTCAACATTTGCAAAATTAAGTTTCTTGGGTCTTA 731
Db 779 ATGAGCGTTTAGGCATGTTTACATCCAGCACTGTGAAGAGCTGCCAGTGGTGGCTG 838
Qy 732 TTAATGAAGTATTTAGGTAAAGCTACTGCTATTTTATTTTACCAGCAAGAGTAAAGCTT 791
Db 839 CTGATGAATACCTGGGCAATGCGCCGCAATCTCTCTCTGCTGATGAGGGGAAACTA 898
Qy 792 CAACATTTAGAGATGAGTTGACTCATGACATTAATTAATAATTTTATAGAGAACGAGGAT 851
Db 899 CAGCACCTGGAAATGAACCTCACCCAGGATATCATCCCAAGTTCCTGGAAATGAAGAC 958
Qy 852 CGTCTGAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTTAAAA 911
Db 959 AGAAGGTCTCCAGCTTACATTTTACCCAAACTGTCCATTTACTGGAACCTATGATCTGAAG 1018
Qy 912 TCTGTTTATAGCCAGTTAGGATATTAACAAAGTTTCTTAACGGTCCGATTTGAGTGGT 971
Db 1019 AGCGTCTCTGGGTCAACTGGGCACTACTAAGTCTTCAGCAATGGGCTGACCTCTCCGG 1078
Qy 972 GTTACTCAAGAGCTCCATTTAAATTTGAGTAAAGCTGTTCAACAAAGCCGTCTTAACTATT 1031
Db 1079 GTCAAGAGGAGGACCCCTGAAAGCTCTCCAGGCGGTGATTAAGGCTGTGCTGACCATC 1138
Qy 1032 GATGAAAGGTTACCGAGCGCGCGCTATGTTCTGGAAGCTATTTCCAATGAGCAAT 1091
Db 1139 GACGAGAAAGGACTGAAGCTGCTGGGCCATGTTTATAGAGCCATACCATGCTATC 1198
Qy 1092 CCACAGAGCTTAAATTTAATAAACCAATTCGTTTCTGATGATCGAGCAGAACACTAAA 1151
Db 1199 CCCCCGAGGTCAAGTTCACAAACCCCTTTGCTCTTATGATTGAACAATAATACCAAG 1258
Qy 1152 AGCCCATTTTATGGGTGAAGTTGTCAACCCCACTCAGAA 1192
Db 1259 TCTCCCTCTTCATGGGAAAGTGGTGAATCCACCCCAAAA 1299
```

Search completed: December 6, 2002, 22:41:39
Job time : 2697 secs